

---

This is an electronic reprint of the original article.  
This reprint may differ from the original in pagination and typographic detail.

Author(s): Hakala, Jussi H. & Oittinen, Pirkko & Häkkinen, Jukka P.  
Title: Depth Artifacts Caused by Spatial Interlacing in Stereoscopic 3D Displays  
Year: 2015  
Version: Post print

**Please cite the original version:**

Hakala, Jussi H. & Oittinen, Pirkko & Häkkinen, Jukka P. 2015. Depth Artifacts Caused by Spatial Interlacing in Stereoscopic 3D Displays. ACM Transactions on Applied Perception. Volume 12, Issue 1. 1-13. ISSN 1544-3558 (printed). DOI: 10.1145/2699266.

Rights: © 2015 Association for Computing Machinery (ACM). This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in ACM Transactions on Applied Perception. Volume 12, Issue 1. 1-13. <http://doi.acm.org/10.1145/2699266>.

---

All material supplied via Aaltodoc is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

# Depth Artifacts Caused by Spatial Interlacing in Stereoscopic 3D Displays\*

Jussi H. Hakala<sup>†1</sup>, Pirkko Oittinen<sup>1</sup>, and Jukka P. Häkkinen<sup>2</sup>

<sup>1</sup>Aalto University

<sup>2</sup>University of Helsinki

## Abstract

Most spatially interlacing stereoscopic 3D displays display odd and even rows of an image to either the left or right eye of the viewer. The visual system then fuses the interlaced image into a single percept. This row-based interlacing creates a small vertical disparity between the images; however, interlacing may also induce horizontal disparities, thus generating depth artifacts. Whether people perceive the depth artifacts, and if so, what is the magnitude of the artifacts, are unknown. In this study, we hypothesized and tested if people perceive interlaced edges on different depth levels. We tested oblique edge orientations ranging from  $2^\circ$  to  $32^\circ$  and pixel sizes ranging from 16 to 79 arcsec of visual angle in a depth probe experiment. Five participants viewed the visual stimuli through a stereoscope under three viewing conditions: non-interlaced, interlaced, and row-averaged (i.e., where even and odd rows are averaged). Our results indicated that people perceive depth artifacts when viewing interlaced stereoscopic images and that these depth artifacts increase with pixel size and decrease with edge orientation angle. A pixel size of 32 arcsec of visual angle still evoked depth percepts, whereas 16 arcsec did not. Row-averaging images effectively eliminated these depth artifacts. These findings have implications for display design, content production, image quality studies, and stereoscopic games and software.

## 1 Introduction

In spatially interlacing stereoscopic displays, the right and left eyes of the viewer see half of the total number of pixels. A polarizing filter in front of the display is often used to direct the appropriate image data to each eye so that odd and even pixel rows have a different polarization. A viewer then views the display through polarizing glasses that direct only the appropriately polarized light to the intended eye, inviting examination of what the viewer perceives as a result of this process. Because each eye sees a different image, horizontal differences in the images produce a perception of three-dimensionality. From a resolution perspective, it is logical to argue that the resolution of the image is halved because each eye sees only half of the total number of pixels. However, the left and right eyes receive different information, and thus, their combined information content equals that of a full-resolution image. Figures 1a-1d illustrate the spatial interlacing method. In Figure 1c, alternating rows of the original image pair are interlaced; in Figure 1d, even and odd rows are averaged before being interlaced. This study demonstrates that the visual system matches the interlaced pixel rows based on local features, which cause depth artifacts. Furthermore, we investigate the use of row-averaging and small pixel sizes to neutralize this phenomenon. In this article, the term *interlacing* refers to spatial row-interlacing, which is commonly achieved by means of polarization and the use of a patterned retarder (Hong et al., 2011). The spatially interlacing displays with a passive polarizing filter are also termed passive 3D displays to distinguish them from temporally interlacing displays that separate the left and right images via active shutter glasses.

---

\*© ACM, 2015. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in ACM Transactions on Applied Perception (TAP), Vol 12, Issue 1, (March 2015) <http://doi.acm.org/10.1145/2699266>

<sup>†</sup>Corresponding author, [jussi.h.hakala@aalto.fi](mailto:jussi.h.hakala@aalto.fi).

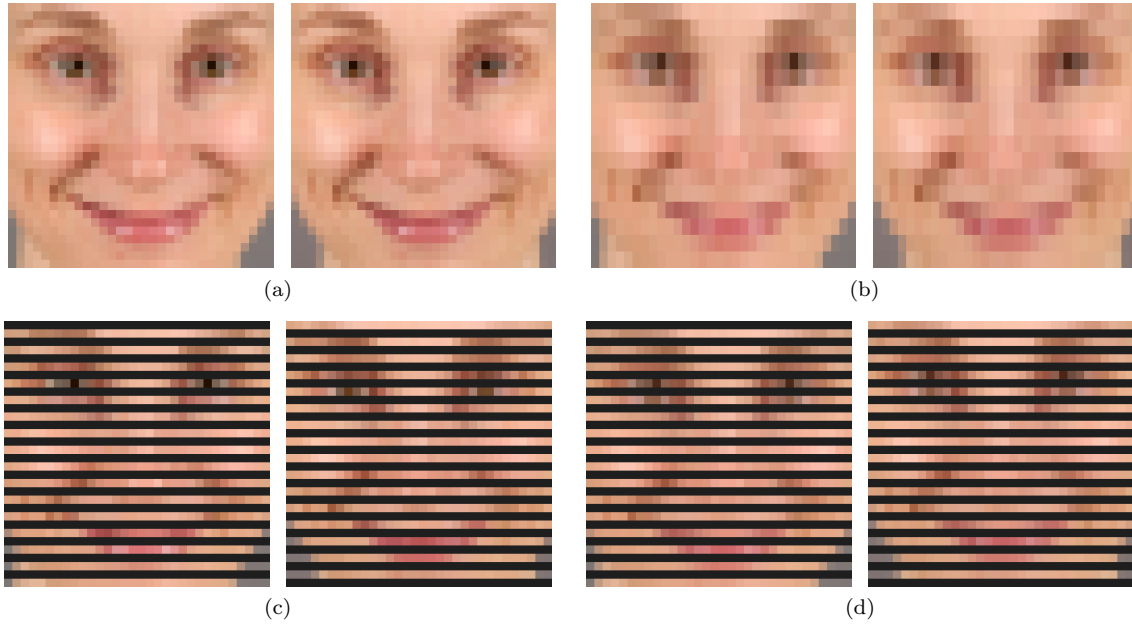


Figure 1: Full-resolution (a) and row-averaged (b) stereoscopic images without interlacing, and (c) and (d) respectively with interlacing, arranged for parallel viewing. In the full-resolution interlaced image (c), the pupils and highlights of the lips are visible in only the left image, whereas row-averaging in (d) blends the details between the interlaced left and right images.

Kelley (2011) compared the perceived resolution in spatially and temporally interlacing stereoscopic displays based on his personal observations. He reported that black interlacing rows are visible at close distances and that the content has a vertical disparity as the alternating pixel rows are fused together. To eliminate this problem, he suggested viewing the display at an adequate distance, where the angular pixel size is below an individual’s ability to resolve detail, and estimated that 60 arcsec is a sufficiently small pixel size. One guideline (Banks et al., 2012) for creating content for interlacing displays suggests that cameras should have a vertical offset in the content creation phase to counteract the one-pixel vertical disparity created by interlacing. This suggestion appears to imply that the visual system fuses the interlaced images as a whole, instead of fusing together the content features of the rows from the left and right images.

Researchers have aimed to determine ways in which interlacing affects the perceived resolution and image quality. Kim and Banks (2012) studied the effects of modifying the interlacing method and viewing distance in a letter recognition experiment. They found that temporal interlacing outperformed spatial interlacing at viewing distances that resulted in Nyquist frequencies of 14 and 28 cycles per degree; however, there was no statistically significant difference between the interlacing methods at a Nyquist frequency of 56 cycles per degree. A Nyquist frequency of 56 cycles per degree corresponds to an angular pixel size of 32 arcsec. The authors also compared monocular and binocular viewing conditions and found no evidence to support the hypothesis that the visual system can fuse interlaced images into a percept where the binocularly perceived resolution surpasses the monocularly perceived resolution. They assume that the eyes create a small vertical vergence movement so that the content rows overlap.

In another study, Yun et al. (2013) used a grating pattern to study contrast sensitivity in spatially and temporally interlacing 3D displays. Their results indicated that the participants could discern Nyquist frequency gratings on a temporally interlacing display but not on a spatially interlacing display. They used a 1920×1080 pixel 47” passive 3D television positioned at a viewing distance of three times the height of the display, which corresponds to an angular pixel size of 64 arcsec. The authors also noted that images displayed on spatially interlacing displays could exhibit jagged edges. Jagged edges were also reported in another study (Park and Choi, 2013) in which participants compared the resolution of temporally and spatially interlacing

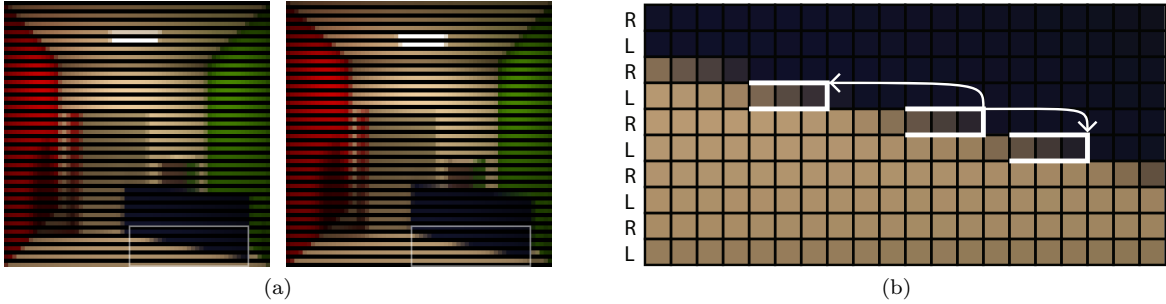


Figure 2: a) An interlaced stereoscopic image pair arranged for parallel viewing. The oblique edge of the dark object is highlighted with a white rectangle. b) The interlaced oblique edge displayed such that the odd pixel rows are intended for the right eye (R) and even pixel rows are intended for the left eye (L). The white edges highlight the segment endpoints and the arrows point to the closest matches in the left eye image for the specific feature in the right eye image.

displays, where interlacing conditions were simulated using a stereoscope. The majority of the participants in that experiment preferred the temporally interlaced condition because they experienced jagged edges in the spatially interlaced condition even though they could not see the black horizontal interlacing rows.

Park et al. (2014) used a two-alternative forced choice experiment to study the effect of the interlacing method on the perceived resolution of natural images. In addition to the temporal and spatial interlacing methods, they also included row-averaged non-interlaced and row-averaged interlaced methods, in which the vertical resolution was half of the original. In row-averaged content, each row is an average of the original odd and even rows, as Figure 1b illustrates. Their study found that the 75% detection threshold for distinguishing the interlaced image from the reference image was reached at a pixel size of 54 arcsec. However, the authors noted that some of their stimuli lacked high-spatial-frequency components and that high-frequency-component imagery requires smaller angular pixel sizes.

Studies on the effects of interlacing have focused on detectability and subjective image quality; however, the issue of how the visual system combines the interlaced views of each eye into a single three-dimensional percept has not been adequately investigated. The binocular correspondence problem addresses how the visual system finds corresponding features within the views of the left and right eyes. For a review on binocular correspondence, see Howard and Rogers (2012, pp. 182-209). The matched features that are close to corresponding areas on each retina are fused. However, interlacing breaks the image features into strips that are separated by black interlacing rows, creating new features that can be matched. Thus, the magnitude of the minimum vertical disparity of the corresponding points between the images of the left and right eyes is equal to the pixel size of the display.

In the case of an oblique edge (e.g., the white rectangle highlighted in Figure 2a), interlacing divides the edges into segments with endpoints; the visual system can then use these segments for matching (van Ee and Schor, 2000). However, the purpose of interlacing is not to introduce new features for matching. If we assume that the visual system fuses the local features, the matched features of an oblique edge will have a horizontal disparity, as shown in Figure 2b. Good matching features are found equally on both sides of the edge, which makes the percept ambiguous. The repeating pattern created by interlacing produces an image that is similar to the wallpaper illusion (Brewster, 1844; Vlaskamp et al., 2013), in which a repeating pattern creates a perception of depth. The nearest neighbor rule (Arditi et al., 1981) states that the visual system fuses the closest matching features to minimize the absolute disparity; however, this rule applies only in the absence of other images. According to the nearest disparity rule (Mitchison and McKee, 1987; Zhang et al., 2001), visual matches that minimize the local disparity relative to nearby unambiguous objects are weighted more heavily. Thus, the perceived depth of the dark edge in Figure 2b is a combination of the intended actual disparity of the edge in the non-interlaced image and the unintended depth artifact caused by interlacing, the magnitude of which depends on the segment length and the depth of the surrounding features. The unintended depth artifacts may interfere with depth judgments, degrade the image quality, cause eye discomfort, and cause eyestrain in extreme cases.

In this paper, we report the results of a depth probe experiment that uses oblique edge stimuli to determine whether people perceive depth artifacts on spatially interlacing stereoscopic displays. Based on the aforementioned research, we hypothesized that people perceive depth artifacts on spatially interlacing stereoscopic displays depending on the pixel size and edge orientation. Our secondary objective is to determine whether depth artifacts can be avoided by row-averaging image content or by using a small pixel size, and if so, to determine which pixel size is adequate to eliminate depth artifacts. To answer these research questions, we conducted an experiment in which the participants moved a probe in depth to the same depth level with a stimulus. We used oblique edges in various orientations as the stimuli in our experiment because edges are the basic features that form contours in natural scenes and images (e.g., Geisler, 2008). To simplify the experiment, we examined cases in which the relative and absolute disparities have a common minimum, i.e., cases in which the edge or nearby unambiguous objects have no inherent horizontal disparity. Our results clearly indicated that people perceive depth artifacts on spatially interlacing stereoscopic displays. These artifacts are visible at a pixel size of 32 arcsec but not at 16 arcsec. The participants did not perceive depth artifacts when row-averaging was utilized.

## 2 Methods

### 2.1 Participants

One woman and four men (JHH, RK, MH, KA, and JPH) participated in the experiment. The participants were between 26 and 44 years of age; three of them wore corrective eyewear. We tested the participants' visual acuity at a distance of 3 m using the Lea Numbers test (Lea Test Ltd., Helsinki, Finland); all participants had a visual acuity of 0.8–1.25 (20/25–20/16) in both eyes, which were measured separately. For stereoacuity screening, we used the TNO test; all participants scored 60 arcsec or better. Near phoria was tested using the Maddox Wing test, and far phoria was tested at 1.6 m using a Maddox Wing pattern that was displayed on a computer screen and viewed through a stereoscope. JHH had 8 prism diopters (PD) of exophoria, whereas JPH had 11 PD of esophoria, measured with the Maddox Wing test at a distance of 30 cm. At a distance of 1.6 m, JHH had 3 PD of exophoria and JPH had 4 PD of esophoria. All of the remaining participants had less than 2 PD of heterophoria, and all participants in this study had less than 1 PD of vertical phoria. Vertical and horizontal fixation disparities were measured using a nonius line stimulus with central and peripheral fixation locks (Ukwade, 2000). JHH, MH, and JPH had 3, 8, and 10 arcmin of horizontal fixation disparity and 3, 1, and 3 arcmin of vertical fixation disparity, respectively. The fixation disparities of the remaining participants were smaller than 1 arcmin. The participants' interpupillary distances were between 60 and 66 mm. Two of the participants were expert evaluators and three were naïve evaluators. JHH and JPH, who are experienced viewers of stereoscopic content, are also the authors of this paper. MH and KA had some experience with stereoscopic content; RK had minimal experience with stereoscopic content.

### 2.2 Stimuli

The stimuli were computer-generated oblique edges presented on a 204 ppi 22.2" ViewSonic VP2290b display, viewed through a mirror stereoscope at a distance of 1.6 m. Figure 3 shows an example of the stimuli. The following four independent variables were defined as the stimulus conditions: edge orientation, interlacing method, pixel size, and edge interpolation technique. The orientation was measured from the horizontal position and varied between  $2^\circ$  and  $32^\circ$  in two-decibel intervals, resulting in seven ascending orientations:  $2.0^\circ$ ,  $3.2^\circ$ ,  $5.0^\circ$ ,  $7.9^\circ$ ,  $12.6^\circ$ ,  $20.0^\circ$ , and  $31.6^\circ$ . The same orientation magnitudes were used for descending edges, and thus, fourteen orientations were tested. In addition to the native pixel size of the display, we used squares that were 2, 3, 4, and 5 pixels wide to simulate larger pixel sizes. At the viewing distance in our experiment, one native display pixel subtended 16 arcsec, resulting in the following angular sizes: 16, 32, 48, 63, and 79 arcsec. The interlacing method conditions were as follows: non-interlaced, interlaced, and row-averaged. In the non-interlaced condition, both eyes viewed all pixels of the edge. In the interlaced condition, the left eye viewed odd pixel rows, and the right eye viewed even pixel rows; simulated pixel rows were viewed in the larger pixel simulations. In the row-averaged condition, the values of the first and second pixel rows were averaged to form the first content row, the third and fourth pixel rows were averaged to form the second content row, and so on, after which the stimulus was interlaced. We used the nearest neighbor



Figure 3: A sample of the left and right eye images that were used as stimuli in the experiment under the interlaced condition using nearest neighbor interpolation, displayed for parallel viewing. The interlacing rows are white and the pixel size is exaggerated for illustrative purposes.

and bilinear techniques for the interpolation of the edge. We applied the interpolation techniques prior to image scaling, which was performed to simulate the larger pixel sizes. Two repeats resulted in a total of 840 trials for each participant.

The area above the edge was darker than the area below the edge. The upper half of the non-interlaced stimulus had a luminance of  $15.2 \text{ cd/m}^2$ , and the lower half had a luminance of  $60.6 \text{ cd/m}^2$ , as measured through the stereoscope with a spectroradiometer (Photo Research SpectraScan PR-670). Under the interlaced conditions, the black interlacing lines had a luminance of  $0.4 \text{ cd/m}^2$ . To keep the overall luminance equal across all of the conditions, we increased the luminance of the edge lines that were shown under the interlaced conditions. Thus, the Michelson contrast was 0.60 between the upper and lower halves under all conditions. To facilitate fusion, we included fusion locks within the stimuli; these locks were vertical black bars located in each corner, measuring 4 arcmin wide and 24 arcmin high, situated  $4^\circ$  diagonally from the center of the stimulus. The edge and fusion locks had no horizontal disparity between the left and right views. Between each trial, the participants viewed a binocular noise image with a fixation cross in the center for 1 s. The participants viewed the stimuli in a dark room through  $8^\circ$  apertures in the stereoscope. The probe was a 24 arcmin black square situated  $1^\circ$  above the center of the edge; participants adjusted the depth of this probe.

### 2.3 Procedure

We used a full-factorial experiment design with a method of adjustment and no time constraints. The participants performed the two experimental blocks on separate days. The first block included the nearest neighbor interpolation condition, and the second block included the bilinear interpolation condition. We instructed the participants to look at the edge in the middle of the stimulus and move the probe to the same depth level at which they perceived the edge to lie. Moving the probe changed the binocular disparity of the probe in a symmetrical manner. At the beginning of each trial, the disparity of the probe was set to a random value between -15 and 15 arcmin to minimize hysteresis effects. During the trial, the movement of the probe was limited to between -53 and 53 arcmin to avoid excessive disparities that would be difficult to fuse. The participants controlled the experiment by pressing a key on the keyboard to start a new trial and make the probe visible. The participants then adjusted the probe disparity to the appropriate depth level by moving the computer mouse back and forth before pressing a key to lock their response. At the beginning of each block, the participants practiced the task under 12 trials. The stimulus order was randomized within each repeat cycle; the participants saw each condition once in a randomized order and then all conditions again in a randomized order.

## 3 Results

Figure 4 presents the main results. The results indicated that the participants perceived the edges in the interlaced condition at different depths, as illustrated in the middle column graphs of Figure 4. Additionally, large orientation angles and small pixel sizes diminished the perceived depth. Furthermore, the participants

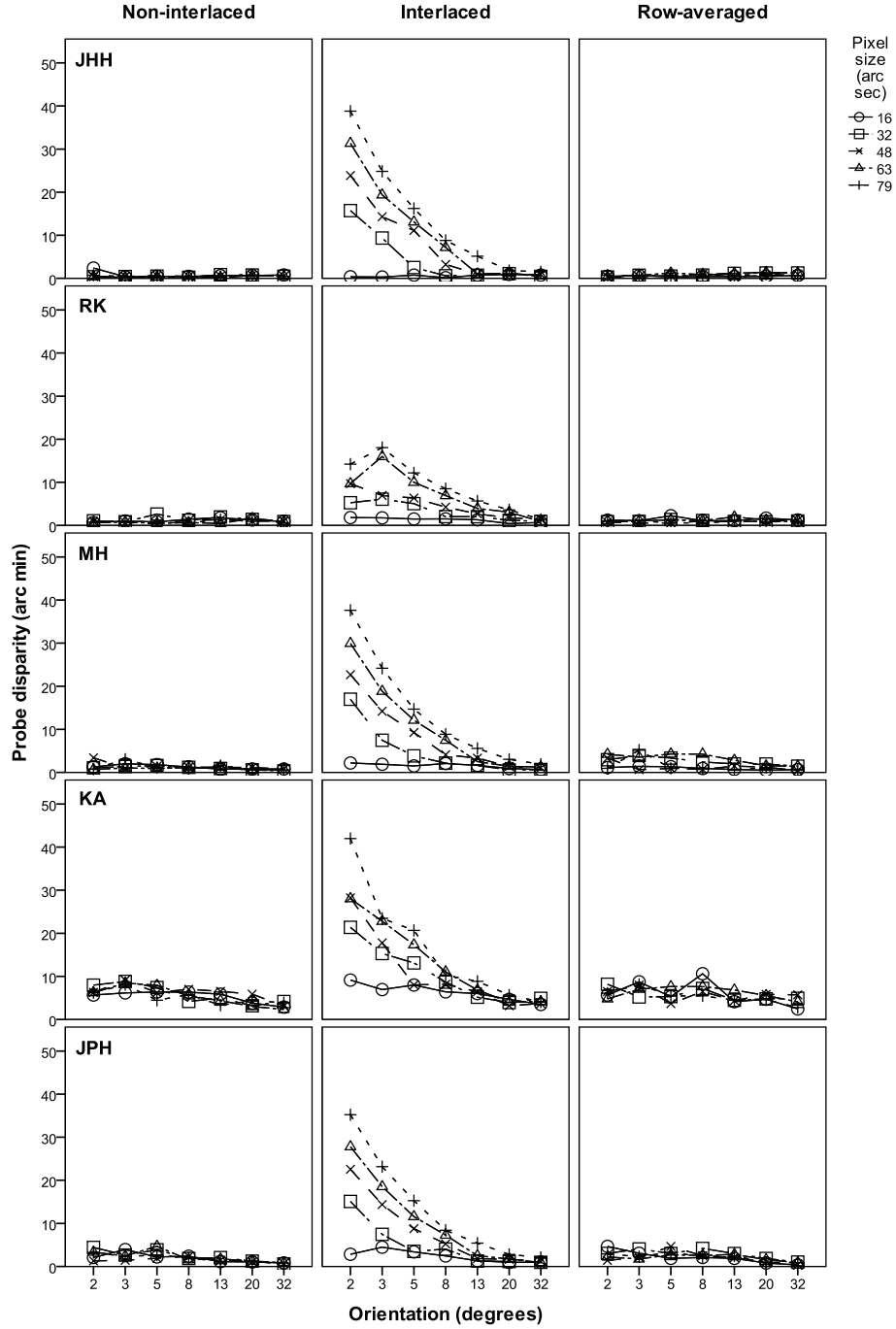


Figure 4: Mean probe disparities for each participant under the non-interlaced, interlaced, and row-averaged conditions. Each participant's data are depicted on a separate row; each interlacing condition is depicted on a separate column. The x-axes represent the absolute edge orientation angles, and the y-axes represent the absolute probe disparities. Each data point is the mean of eight trials.

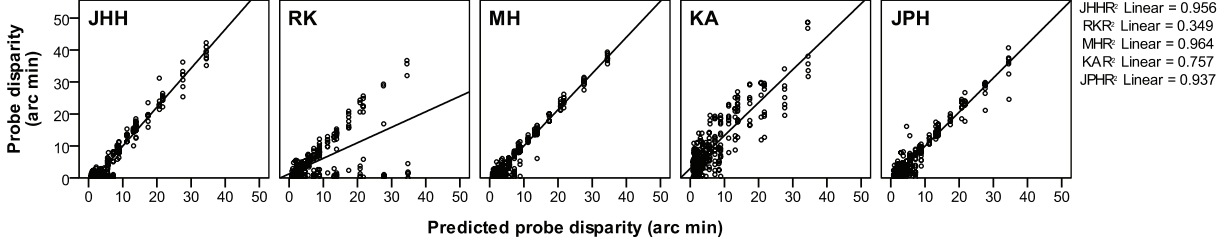


Figure 5: Linear regression analysis conducted separately for each participant. The x-axes represent the values predicted by each regression model; the y-axes represent the participants’ probe disparities.

perceived all non-interlaced and row-averaged stimuli approximately at the display plane (the first and last columns of Figure 4). For our data analyses, we separated the probe disparity and the direction of the disparity. The probe disparity is the magnitude of the horizontal disparity of the probe that is required to perceive the probe at the same depth as the edge. The direction of the disparity, which can either be crossed or uncrossed, represents whether the participant perceived the edge as being in front of the display or behind the display, respectively. We also studied the edge orientation magnitude and sign separately. The probe disparity was dependent on the pixel size and orientation magnitude, as hypothesized. A repeated-measures analysis of variance was performed with the orientation magnitude, orientation sign, pixel size, interlacing method, and interpolation technique as the variables. Significant ( $p < .01$ ) main effects for orientation magnitude ( $F(1.3, 5.1) = 31.2, \omega^2 = .84$ ), pixel size ( $F(1.3, 5.1) = 72.6, \omega^2 = .87$ ), and interlacing method ( $F(1.1, 4.3) = 146.9, \omega^2 = .95$ ) were verified with Greenhouse-Geisser corrected estimates. The interpolation technique and orientation sign had no significant effects on the probe disparity. A three-way interaction analysis between the orientation magnitude, pixel size, and interlacing method also indicated a significant effect ( $F(2.4, 9.4) = 20.0, \omega^2 = .74, p < .001$ ). Whereas the interlaced condition differed significantly from the non-interlaced condition ( $F(1, 4) = 142.4, r = .99, p < .001$ ), the difference between the row-averaged condition and the non-interlaced condition was small and did not quite reach statistical significance ( $F(1, 4) = 5.6, r = .76, p = .076$ ). The sample standard deviation was slightly higher in the row-averaged condition ( $s = 3.5$ ) compared with the non-interlaced condition ( $s = 3.0$ ).

All orientation magnitudes that were used in the experiment created significant differences in the probe disparity when compared with the adjacent orientations. The difference between the interlaced and non-interlaced conditions was pronounced at small orientation magnitudes. At the steepest orientation magnitude ( $31.6^\circ$ ), the difference between the interlaced and non-interlaced conditions was significant ( $F(1, 4) = 20.2, r = .91, p = .011$ ); however, at the steepest orientation, the mean of the interlaced condition was only 0.44 arcmin higher than the mean of the non-interlaced condition. Furthermore, this difference was no longer statistically significant when we excluded the largest pixel size from the analysis. In addition, all size conditions differed significantly from their adjacent size conditions; larger pixel sizes created larger probe disparities. When we analyzed the smallest pixel size (16 arcsec), we found no statistically significant differences between the interlaced and non-interlaced conditions. When we analyzed the next larger pixel size (32 arcsec), the difference between the interlaced and non-interlaced conditions was significant ( $F(1, 4) = 42.2, r = .96, p < .01$ ); statistical significance was observed for all of the larger pixel size conditions.

To test the relationship between the edge segment length and probe disparity, we built a linear regression model for each participant based on the edge segment length in each interlaced condition. The edge segment length depends on the pixel size and orientation magnitude such that segment length  $l = \text{size} / \tan(\text{orientation})$ . We excluded data points from the analysis if they had residual values that exceeded three standard deviations. Figure 5 shows the result of the regression analysis. The goodness of fit was very good ( $R^2 > .90$ ) for JHH, MH, and JPH. However, the data from KA were considerably noisier than the data from the other participants; therefore, the goodness of fit was poorer ( $R^2 = .76$ ). RK provided near-zero evaluations for several stimuli that had long edge segment lengths; thus, the goodness of fit was poor compared with those of the other participants ( $R^2 = .35$ ).

To analyze whether the participants saw the edge in front of or behind the display plane under the interlaced condition, we divided the probe disparities into the following three location categories: in front,



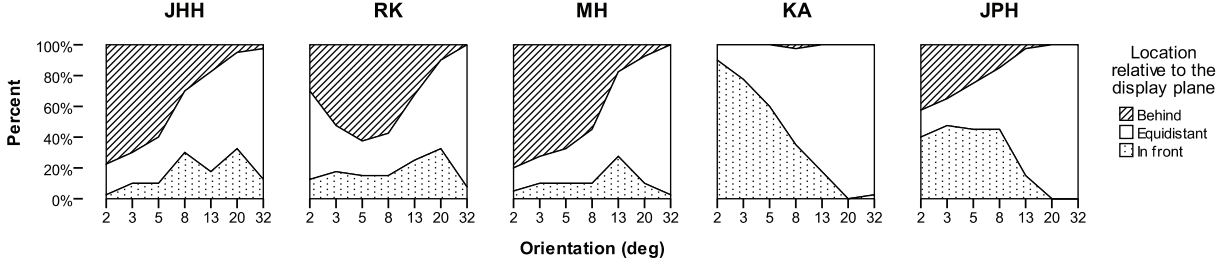


Figure 6: The perceived location of the edge relative to the display plane for each participant. The x-axes indicate the orientation angle, and the y-axes indicate the percentage of trials belonging to each category. Dashed regions represent cases in which the participant placed the probe behind the display plane, dotted regions represent placement in front of the display plane, and white regions represent placement equidistant to the display plane.

equidistant, and behind. We defined equidistance to the display plane as a probe disparity value that does not exceed the mean of the non-interlaced condition response of the participant, which was near zero disparity for most participants, by more than one standard deviation. Figure 6 shows the ratios of the perceived location as a function of the orientation for each participant. JHH, RK, and MH perceived the depths under the interlaced condition mostly behind the display plane, i.e., the probe had a positive disparity. By contrast, KA perceived nearly all depths in front of the display plane. JPH perceived a majority of the edges in front of the display plane but also perceived some of the edges behind the display plane. The randomized probe disparity at the beginning of the trial had no effect on the direction of the perceived disparity. Furthermore, there was no consistency regarding the relationship between the direction of the perceived disparity and the vision test results (e.g., heterophoria or fixation disparity). The mean trial durations were 8, 23, 9, 7, and 8 s for JHH, RK, MH, KA, and JPH, respectively.

## 4 Discussion

Our objective was to determine whether spatial interlacing causes depth artifacts. The results of our study clearly indicate that depth artifacts are present in oblique edges that are viewed on a simulated spatially interlacing stereoscopic display. The participants’ depth estimates were consistent with the disparity created by the segment length of the oblique edge. Our secondary objective was to determine whether depth artifacts can be eliminated by row-averaging image content or using a small pixel size, and if so, to determine which pixel size is sufficiently small to eliminate depth artifacts. Row-averaging eliminated depth artifacts effectively at all pixel sizes. We tested five pixel sizes between 16 and 79 arcsec and found that only the smallest pixel size (16 arcsec) did not generate any depth artifacts.

Our findings indicate that spatially interlacing stereoscopic displays should either use content that is row-averaged or have a pixel size that is considerably smaller than previously suggested in other published studies. As mentioned above, earlier research has suggested that pixel sizes 60 arcsec (Kelley, 2011), 32 arcsec (Kim and Banks, 2012), and 56 arcsec Park et al. (2014) are sufficiently small for displaying content on interlacing stereoscopic displays. To prevent the visual system from resolving the content rows and black interstitial rows, the pixel size should be smaller than the visual acuity, which is 60 arcsec for individuals with 20/20 vision. However, our results showed that people could still perceive depth artifacts even though they were unable to resolve the pixel rows. A simple explanation for this discrepancy is related to the way the image is interlaced. The alternating content and interlacing rows halve the vertical resolution. At pixel size of 32 arcsec, the spatial interlacing frequency is 56 cycles per degree. The participants could not resolve the interlacing rows from the content rows at such a high frequency, but they could resolve the two content rows, which were 64 arcsec apart. Thus, the visual system can use this information in disparity processing. The visibility of edge jaggedness in computer graphics depends on the orientation of the edge. The “jags” are least visible near the horizontal and vertical orientations but also at the 45° orientation (Naiman, 1998). In stereoscopic interlacing displays, orientations of 32° and above do not create depth artifacts unless the pixel size is extremely large. However, we did not consider the visibility of jaggedness in this study.

The vertical disparity created by interlacing is a result of the optical design used to interlace the images on the display. Thus, balancing the vertical disparity using a camera arrangement as suggested earlier (Banks et al., 2012) seems infeasible. Furthermore, offsetting the cameras would change the perspective, further distorting the view of the scene. Instead, matching pixel rows from the left and right camera images could be used (i.e., only odd rows from both images) to eliminate depth artifacts. However, in practice, optical distortions and the difficulty of offsetting physical cameras to such accuracy make this option challenging. In computer graphics, a precise offset could be implemented; however, due to the jaggedness and moiré that is caused by subsampling, row-averaging should be used instead. Row-averaging can be achieved by capturing the scene at half of the vertical resolution to conserve processing or rendering time for real-time applications.

When the visual system finds matching images, it can use these linked images to adjust the rotation of the eyes to find the global maximum correspondence. The depth latency that is caused by the process of finding this global maximum correspondence (Howard and Rogers, 2012, p. 184) may explain the long trial durations observed in our study. The linked local features that have a horizontal disparity can induce horizontal vergence; furthermore, vertical disparity can induce vertical vergence. Earlier research has shown that stimuli at 45 arcmin evoke vertical vergence that is nearly as strong as the vertical vergence evoked by 65° stimuli (Howard et al., 2000); therefore, we can safely assume that the 8° stimuli used in our experiment are sufficiently large to evoke vertical vergence. In an experiment with horizontal line stimuli, Duwaer and van den Brink (1981) found that a disparity as small as 0.4 arcmin can induce vergence eye movements in foveal stimuli. Thus, row-interlacing stereoscopic displays likely induce vertical vergence eye movements that compensate for the vertical disparity. In our experiment, the participants viewed the stimuli through stationary apertures. Therefore, the participants’ visual system had a strong reference without vertical disparity. In this sense, our experimental condition mimicked a natural viewing condition where the environment surrounding the interlacing display is visible. Because the gain of vertical vergence is considerably weaker in peripheral vision than in the center (Howard et al., 2000), the interlacing pattern of the display likely triggers vertical vergence in our experiment as well as in a regular viewing situation.

In natural stereoscopic content, other objects can affect the fusion of the edge via a disparity gradient (Burt and Julesz, 1980) or stereo capture (Ramachandran and Cavanagh, 1985). Furthermore, the matching of ambiguous features is affected by nearby objects and defined by the nearest disparity rule (Mitchison and McKee, 1987). In this experiment, the unambiguous disparities of the nearby objects, such as the fusion locks, draw the disparity of the ambiguous edge toward them. The stimuli in our study were limited to 8° of viewing angle, which meant that the edge had an unambiguous matching feature on the left and right borders of the stimulus. Research has shown that when viewers are faced with ambiguous matches, unambiguous matches can force one of the discrete ambiguous matches that is available; alternatively, in special cases, the visual system interpolates the depth based on the endpoint disparities (Mitchison and McKee, 1985). In our experiment, two participants commented that some of the edges appeared slanted or curved in depth, which could indicate the effect of these endpoints. None of the participants reported stereo capture, i.e. that moving the probe would have changed the perceived depth of the edge. Nevertheless, some stereo capture probably occurred because of the closeness of the probe to the edge. The vertical disparity that is created by the matching of different pixel rows can also contribute to the perceived slant of the edge (Backus et al., 1999). Depth artifacts caused by interlacing are relative to the content depth; in our experiment, stimuli were located on the display plane. If the content has an inherent disparity, the nearest match may not necessarily be on an adjacent row. Instead, the nearest match may be, for example, three to five lines above or below the feature being matched; this effect would further increase the local vertical disparity.

RK provided near-zero responses to stimuli that had large horizontal disparities. Furthermore, RK’s trials lasted significantly longer than the other participants’ trials. Although recent results indicate that the influence of experience on depth discrimination, depth matching, and fusion limits is small (Stransky et al., 2014), a learning process related to perceiving ambiguous stereoscopic images was previously reported (Julesz, 1960). We suspect that RK’s inexperience with stereoscopic images contributed to the longer time required to achieve fusion. Another explanation could be individual differences in Panum’s fusional area size. If the horizontal stimulus disparity was above the limit for horizontal disparity, the edge would appear as a double image. In such a case, the participant would respond based on the successfully fused images in the stimulus, such as the fusion locks, which had zero disparity. The bimodality of RK’s responses, which is evident in Figure 5, supports this theory.

In this study, we held the disparity of the oblique edge and fusion locks, which surrounded the edge,

at zero, so that the absolute and relative disparities had a common minimum. The participants perceived the edge either behind or in front of the display. This preference varied between the individuals in our study. Future studies should address this issue and determine why people exhibit a strong preference for negative parallax or positive parallax or no strong preference in either direction. In our sample, we did not obtain an explanation for the inter-individual differences from the heterophoria or fixation disparity test results. Furthermore, evaluation of the effects of interlacing on the perception of dynamic content is also an important subject for future studies. The content that is displayed on stereoscopic 3D displays is often dynamic, while the stimuli in our experiment were stationary. For example, for an edge moving vertically, interlacing can cause flickering as each edge segment alternates between the left and right views. Rotation of the edge changes the edge orientation and thus also the perceived depth of the edge as is evident from our results.

The results indicate that in spatially interlacing stereoscopic displays, black interstitial lines divide the underlying image features into smaller features, which the visual system uses in disparity processing. In addition to the misperceptions (Held and Banks, 2008) and artifacts (Boev et al., 2009) reported earlier, the disparity artifacts described in this paper should be considered when designing stereoscopic displays, studying stereoscopic image quality, creating content for stereoscopic displays, and developing software for stereoscopic games and media playback.

## 5 Conclusions

In this paper, we showed that people perceive depth artifacts on spatially interlacing stereoscopic displays. In a depth probe experiment with oblique edges as stimuli, participants consistently reported depth artifacts unless the underlying image was row-averaged or the pixel size was as small as 16 arcsec. Previous research has focused on the perceived resolution and image quality on stereoscopic displays and has recommended pixel sizes of 32 or 56 arcsec for spatially interlacing stereoscopic displays. Although a pixel size of 32 arcsec has been found to be sufficiently small to prevent the loss of perceived acuity, our findings indicated that this pixel size can still elicit depth artifacts. Our study increased the body of knowledge on mechanisms through which the visual system fuses the spatially interlaced images into a three-dimensional percept.

To eliminate depth artifacts, stereoscopic content displayed on spatially interlacing displays requires row-averaging, unless the pixel size is extremely small. Based on our results, pixel sizes of 16 arcsec or smaller are appropriate without row-averaging. Thus, on a high-definition display (e.g., 1920×1080 pixels), the viewing distance must be up to 12 times the display height to avoid depth artifacts; however, in many cases, this viewing distance is infeasible. Another method of avoiding depth artifacts is to display row-averaged content, which reduces the vertical resolution by half. Interlacing without row-averaging should be avoided, although it seems advantageous from an information preservation perspective.

This research was supported by Academy of Finland project "Computational Psychology of Experience in Human-Computer Interaction" (Project number 264323).

### \*Bibliography

- Aries Arditi, Lloyd Kaufman, and J. Anthony Movshon. 1981. A simple explanation of the induced size effect. *Vision Research* 21, 6 (1981), 755–764. DOI:[http://dx.doi.org/10.1016/0042-6989\(81\)90173-5](http://dx.doi.org/10.1016/0042-6989(81)90173-5)
- Benjamin T. Backus, Martin S. Banks, Raymond van Ee, and James A. Crowell. 1999. Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research* 39, 6 (1999), 1143–70. DOI:[http://dx.doi.org/10.1016/S0042-6989\(98\)00139-4](http://dx.doi.org/10.1016/S0042-6989(98)00139-4)
- Martin S. Banks, Jenny R. Read, Robert S. Allison, and Simon J. Watt. 2012. Stereoscopy and the Human Visual System. *SMPTE Motion Imaging Journal* 121, 4 (2012), 24–43. DOI:<http://dx.doi.org/10.5594/j18173>
- Atanas Boev, Danilo Hollosi, Atanas Gotchev, and Karen Egiazarian. 2009. Classification and simulation of stereoscopic artifacts in mobile 3DTV content. *Proc. SPIE* 7237 (2009), 72371F–72371F–12. DOI:<http://dx.doi.org/10.1117/12.807185>
- David Brewster. 1844. On the Knowledge of Distance given by Binocular Vision. *Transactions of the Royal Society of Edinburgh* 15, 4 (1844), 663–675. DOI:<http://dx.doi.org/10.1017/S0080456800030246>

- Peter Burt and Bela Julesz. 1980. Modifications of the classical notion of Panum’s fusional area. *Perception* 9, 6 (1980), 671–82. DOI:http://dx.doi.org/10.1068/p090671
- Arne Lex Duwaer and G. van den Brink. 1981. Diplopia thresholds and the initiation of vergence eye-movements. *Vision Research* 21, 12 (1981), 1727–1737. DOI:http://dx.doi.org/10.1016/0042-6989(81)90205-4
- Wilson S. Geisler. 2008. Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology* 59 (2008), 167–92. DOI:http://dx.doi.org/10.1146/annurev.psych.58.110405.085632
- Robert T. Held and Martin S. Banks. 2008. Misperceptions in stereoscopic displays: a vision science perspective. In *Proceedings of the 5th symposium on Applied Perception in Graphics and Visualization*, Vol. 1. 23–32.
- Jisoo Hong, Youngmin Kim, Hee-Jin Choi, Joonku Hahn, Jae-Hyeung Park, Hwi Kim, Sung-Wook Min, Ni Chen, and Byoungcho Lee. 2011. Three-dimensional display technologies of recent interest: principles, status, and issues [Invited]. *Applied Optics* 50, 34 (2011), H87–115. DOI:http://dx.doi.org/10.1364/AO.50.000H87
- Ian P. Howard, Xueping Fang, Robert S. Allison, and James E. Zacher. 2000. Effects of stimulus size and eccentricity on horizontal and vertical vergence. *Experimental Brain Research* 130, 2 (2000), 124–32. DOI:http://dx.doi.org/10.1007/s002210050014
- Ian P. Howard and Brian J. Rogers. 2012. *Perceiving in Depth, Volume 2: Stereoscopic Vision*. Oxford University Press, New York, New York, USA.
- Bela Julesz. 1960. Binocular Depth Perception of Computer-Generated Patterns. *The Bell System Technical Journal* 39, 5 (1960), 1125–1162. DOI:http://dx.doi.org/10.1002/j.1538-7305.1960.tb03954.x
- Edward F. Kelley. 2011. Resolving Resolution. *Information Display* 27, 9 (2011), 18–21.
- Joohwan S. Kim and Martin S. Banks. 2012. 65.1: Effective Spatial Resolution of Temporally and Spatially Interlaced Stereo 3D Televisions. In *SID Symposium Digest of Technical Papers*, Vol. 43. 879–882. DOI:http://dx.doi.org/10.1002/j.2168-0159.2012.tb05927.x
- Graeme J. Mitchison and Suzanne P. McKee. 1985. Interpolation in stereoscopic matching. *Nature* 315, 6018 (1985), 402–404. DOI:http://dx.doi.org/10.1038/315402a0
- Graeme J. Mitchison and Suzanne P. McKee. 1987. The resolution of ambiguous stereoscopic matches by interpolation. *Vision Research* 27, 2 (1987), 285–294. DOI:http://dx.doi.org/10.1016/0042-6989(87)90191-X
- Avi C. Naiman. 1998. Jagged edges: when is filtering needed? *ACM Transactions on Graphics* 17, 4 (1998), 238–258. DOI:http://dx.doi.org/10.1145/293145.293147
- Minyoung Park and Hee-Jin Choi. 2013. Comparison of perceived resolution between the active retarder and patterned retarder 3D displays. In *Digital Holography and Three-Dimensional Imaging*. OSA, Washington, D.C., DM2A.6. DOI:http://dx.doi.org/10.1364/DH.2013.DM2A.6
- Minyoung Park, Joohwan Kim, and Hee-Jin Choi. 2014. Effect of interlacing methods of stereoscopic displays on perceived image quality. *Applied Optics* 53, 3 (2014), 520–527. DOI:http://dx.doi.org/10.1364/AO.53.000520
- Vilayanur S. Ramachandran and Patrick Cavanagh. 1985. Subjective contours capture stereopsis. *Nature* 317, 6037 (1985), 527–530. DOI:http://dx.doi.org/10.1038/317527a0
- Debi Stransky, Laurie M. Wilcox, and Robert S. Allison. 2014. Effects of Long-Term Exposure on Sensitivity and Comfort with Stereoscopic Displays. *ACM Transactions on Applied Perception* 11, 1 (2014), 1–13. DOI:http://dx.doi.org/10.1145/2536810

- Michael T. Ukwade. 2000. Effects of Nonius Line and Fusion Lock. *Optometry and Vision Science* 77, 6 (2000), 309–320. DOI:<http://dx.doi.org/10.1097/00006324-200006000-00011>
- Raymond van Ee and Clifton M. Schor. 2000. Unconstrained stereoscopic matching of lines. *Vision Research* 40, 2 (2000), 151–62. DOI:[http://dx.doi.org/Unconstrained stereoscopic matching of lines](http://dx.doi.org/Unconstrained%20stereoscopic%20matching%20of%20lines)
- Björn N. S. Vlaskamp, Phillip Guan, and Martin S. Banks. 2013. The venetian-blind effect: a preference for zero disparity or zero slant? *Frontiers in Psychology* 4, November (2013), 836.1–9. DOI:<http://dx.doi.org/10.3389/fpsyg.2013.00836>
- Joo Dong Yun, Youngshin Kwak, and Seungjoon Yang. 2013. Evaluation of Perceptual Resolution and Crosstalk in Stereoscopic Displays. *Journal of Display Technology* 9, 2 (2013), 106–111. DOI:<http://dx.doi.org/10.1109/JDT.2012.2228252>
- Zhilei Zhang, Mark Edwards, and Clifton M. Schor. 2001. Spatial interactions minimize relative disparity between adjacent surfaces. *Vision Research* 41, 23 (2001), 2995–3007. DOI:[http://dx.doi.org/10.1016/S0042-6989\(01\)00179-1](http://dx.doi.org/10.1016/S0042-6989(01)00179-1)